

## BIROn - Birkbeck Institutional Research Online

Eimer, Martin (2015) EPS mid-career award 2014: the control of attention in visual search - cognitive and neural mechanisms. *The Quarterly Journal of Experimental Psychology* 68 (12), pp. 2437-2463. ISSN 1747-0218.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/12468/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>  
contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

or alternatively

**EPS Mid-Career Award 2014:**

**The control of attention in visual search - Cognitive and neural mechanisms**

Martin Eimer

Department of Psychological Sciences, Birkbeck College, University of London,  
Malet Street, London WC1E 7HX, UK

Phone: 0044 20 76316358

Fax: 0044 20 76316312

Email: [m.eimer@bbk.ac.uk](mailto:m.eimer@bbk.ac.uk)

**Acknowledgement.** This work was supported by grant ES/L016400/1 from the Economic and Social Research Council (ESRC), UK. I would like to thank the members of my research group, and in particular Anna Grubert, for many helpful comments on previous versions of this article.

**Abstract**

In visual search, observers try to find known target objects among distractors in visual scenes where the location of the targets is uncertain. This review article discusses the attentional processes that are active during search and their neural basis. Four successive phases of visual search are described. During the initial preparatory phase, a representation of the current search goal is activated. Once visual input has arrived, information about the presence of target-matching features is accumulated in parallel across the visual field (guidance). This information is then used to allocate spatial attention to particular objects (selection), before representations of selected objects are activated in visual working memory (recognition). These four phases of attentional control in visual search are characterized both at the cognitive level and at the neural implementation level. It will become clear that search is a continuous process that unfolds in real time. Selective attention in visual search is described as the gradual emergence of spatially specific and temporally sustained biases for representations of task-relevant visual objects in cortical maps.

## 1. Introduction

The visual world is rich and complex. Many visual objects and events are simultaneously received by the visual system, but only some of these are linked to current intentions and action goals. To facilitate adaptive behaviour, these task-relevant objects need to be preferentially processed so that they can be rapidly detected and recognized. Such goal-dependent modulations of visual processes and their effects on perception and action are described as 'selective attention'. Attentional mechanisms affect the perception and recognition of visual objects in different task contexts, and have been investigated with various experimental procedures. Many visual attention experiments have employed spatial cues that inform observers about the likely locations of upcoming target objects (e.g., Posner, Snyder, & Davidson, 1980). Such advance spatial information can be used to move focal spatial attention to particular visual field locations in anticipation of task-relevant objects at these locations, which facilitates the detection and recognition of these objects.

However, there are many other situations where the selective processing of visual input is required, but no precise advance spatial information about the location of task-relevant objects is available. Trying to find the missing car keys in a cluttered office room is a challenge when we cannot remember where we left them a few minutes before. In everyday life, there are many such instances where visual search is required to find a known target object at an unknown location. Although search in the real world will sometimes benefit from contextual information about the likely location of particular objects (e.g., alarm clocks are frequently found on bedside tables; see Henderson, 2003), many lab-based visual search experiments require participants to find specific target objects or features at random and thus entirely unpredictable locations (e.g., Treisman & Gelade, 1980). Under certain conditions, this is very easy: When a target object has a unique visual feature (e.g., when the target is the only red object among green distractors), it will 'pop out' from its surroundings and can be detected rapidly (e.g., Treisman, 1988; Müller, Heller, & Ziegler, 1995). In many other situations, a search target cannot be found on the basis of its perceptual salience alone, and search becomes harder. It is generally believed that the successful detection and recognition of target objects in such situations depends on selective attention. But what role does attention play during visual search, and how do such

attentional processes operate? This article will discuss the cognitive and neural mechanisms that are responsible for our ability to find known target objects at uncertain locations. Some of the ideas developed here have been previously summarized in a brief review article (Eimer, 2014).

Several models of attentional processing such as Feature Integration Theory (e.g., Treisman, 1988) and Guided Search (Wolfe, 1994, 2007) have been developed specifically to explain behavioural performance in various visual search tasks. More general theories of selective attention such as the biased competition account (e.g., Desimone & Duncan, 1995) and the neural theory of visual attention (Bundesen, Habekost, & Kyllingsbaek, 2005) also have direct implications for visual search. Some of the ideas put forward by these models will be considered below. To understand the cognitive and neural mechanisms involved in visual search, it is important to keep in mind that search is a process that unfolds in real time. Even an informal description of this process can readily distinguish different stages that operate sequentially when an observer searches for a known target object at an unknown location. First, the observer has to form an intention to find a specific object, and activate some form of mental representation of this object. Once visual input has arrived, possible target objects have to be localized among other irrelevant objects in the visual field. Next, attention can be selectively focused on one or more of these objects. Finally, particular objects are recognized as targets or nontargets. In this review article, this informal description of four component processes involved in visual search will be used to characterize these search processes at the cognitive level, and to discuss how these cognitive operations are implemented neurally. The four-stage structure of visual search proposed here is illustrated in Figure 1. Preparation, guidance, selection, and recognition are distinguished as separate phases of visual search that jointly contribute to the detection and recognition of particular search targets. Each of these four stages are described at the cognitive level in terms of their functional roles, and at the neural level with respect to the brain processes that implement these particular functions.

In the subsequent sections of this article, each of these four phases of visual search will be considered in turn. During the preparation phase, specific search goals (“attentional templates”) are activated in visual working memory (section 2). Guidance refers to the accumulation of information about goal-relevant features during the initial parallel processing of visual input (section 3). Selection operates through the allocation of focal

attention to possible target objects at particular locations (section 4). Finally, the recognition of selected objects takes place once these objects are encoded into visual working memory (section 5). This four-phase processing model is illustrated in Figure 1 with a box-and-arrow diagram. Such diagrams are frequently used in cognitive psychology to define temporally and functionally discrete stages of information processing. In visual search, such a stage-based model can be heuristically useful to distinguish different aspects of the search process. However, this does not imply that the underlying cognitive and neural mechanisms operate in a strictly modular and discrete fashion. When considering the four phases of visual search and their interactions, it will become clear that search is a continuous process where specific aspects of attentional selectivity emerge gradually in real time.

## **2. Preparatory attentional templates and visual working memory**

Before searching for a particular target object among distractors, observers first have to decide which object or object feature to look for. Next, they have to form a mental representation of the search target, which can be activated in a preparatory fashion before a search display is presented and visual input is processed. The central role of such mental representations for the goal-directed allocation of attention was already highlighted by William James (1890/1981), who referred to “the anticipatory preparation from within of the ideational centers concerned with the objects to which attention is paid” (p.411). According to James, “...the image in the mind *is* the attention; the preperception ... is half of the perception of the looked-for thing” (p.419). In current models of visual attention and visual search, James’ “images in the mind” are described as “attentional templates” (Duncan & Humphreys, 1989; Olivers, Peters, Houtkamp, & Roelfsema, 2011). Such search templates are representations of task-relevant objects or features in visual working memory that are activated while observers prepare for a search task. Attentional templates are set up prior to the presentation of search displays, and then help to direct focal attention towards the location of candidate targets in these displays.

How could such preparatory attentional templates be implemented at the neural level? Because search templates are assumed to be representations in visual working

memory, an answer to this question needs to take into account current views about the neural basis of working memory. It is often assumed that lateral prefrontal cortex plays a critical role in the storage and maintenance of visual information. In line with this hypothesis, neurons in monkey prefrontal cortex show sustained delay activity during the retention period of working memory tasks (e.g., Fuster & Alexander, 1971). However, more recent findings have cast doubt on the hypothesis that prefrontal areas are the primary locus for working memory storage. Human neuroimaging studies have found memory-related delay activity in brain regions outside prefrontal cortex, and in particular in higher-level visual areas in inferior temporal cortex (e.g., Ranganath, Cohen, Dam, & D'Esposito, 2004), suggesting that visual-perceptual cortical regions are also involved in the active short-term maintenance of visual information. This emerging "sensory recruitment" model of visual working memory (Awh & Jonides, 2001; Postle, 2006; D'Esposito, 2007; Harrison & Tong, 2009; Sreenivasan, Curtis, & D'Esposito, 2014) proposes that posterior visual brain areas that are activated during the perception of visual stimuli are also the primary locus for the temporary maintenance of these stimuli in working memory. According to this model, prefrontal areas have more generic top-down control functions, such as regulating access to working memory and maintaining memory representations in an active state through the allocation of focal attention.

If preparatory attentional templates for particular search targets are representations in visual working memory, the sensory recruitment model of working memory predicts that these templates should be implemented by sustained target-specific activation patterns in visual cortex that are similar to the patterns observed when the same target is perceptually processed. This prediction has been investigated in single-unit recording experiments with monkeys and in human functional neuroimaging studies. In these experiments, neural activity was recorded prior to the presentation of search displays while observers prepared for a particular visual search task after being instructed to find a specific target stimulus. The first evidence for a neural correlate of preparatory attentional templates was found in a study by Chelazzi, Duncan, Miller, & Desimone (1998). Monkeys were shown a picture of a specific target object which was then followed after a delay period by a search display that could contain the target object and an irrelevant distractor object or two task-irrelevant objects. The target object had to be retained in working memory during the delay period because the monkeys had to make an eye movement towards this object when it appeared

in the subsequent search display. Neurons in inferotemporal cortex that were selectively activated by the target object during its initial presentation maintained this activation in a sustained fashion prior to the presentation of the search display, suggesting that a representation of the search target was kept active during the delay period. Further evidence for such preparatory “baseline shifts” of neural activity in visual cortex was obtained in human neuroimaging studies. In these studies, where observers prepared to find search targets that were defined by a particular colour or motion, activity in colour- or motion-selective visual brain areas increased during the preparation period prior to the presentation of visual input (e.g., Chawla, Rees, & Friston, 1999; Giesbrecht, Weissmann, Woldorff, & Mangun, 2006). More recent fMRI studies analysed spatially distributed brain activation patterns (multi-voxel pattern analysis, MVPA) to identify preparatory attentional templates in visual cortex. Preparation for a specific target shape elicited a shape-selective pattern of neural activation in lateral occipital cortex during the interval before the target was presented (Stokes, Thompson, Nobre, & Duncan, 2009). Even search for category-defined target objects in natural visual scenes (e.g., people or cars) was found to be preceded by preparatory category-selective activation patterns in visual cortex (Peelen & Kastner, 2011).

These observations suggest that attentional templates are implemented by sustained preparatory changes in the activation pattern of visual areas that are selective to the target feature or object in an upcoming visual search task, and are structurally similar to the activation that is elicited when the same feature or object is visually perceived. This scenario is not only consistent with the sensory recruitment model of working memory, but may also be seen as a 21<sup>st</sup> century neuroscientific confirmation of William James’ 19<sup>th</sup> century suggestion that attentional preparation is based on “images in the mind”. Unfortunately, there is a problem with this parsimonious and intuitively appealing identification of search templates in visual working memory with preparatory baseline shifts of neural activation patterns in visual cortex. To demonstrate that target-selective modulations of neural processing during the preparation for an attentional selection task are the physiological counterpart of search templates, it is necessary to show that such preparatory baseline shifts result in larger sensory responses to target objects once a search display has been presented, and ultimately in the successful detection and identification of these objects (e.g., Driver & Frith, 2000). However, there is so far very limited evidence for



direct links between baseline shifts and the subsequent selective attentional processing of visual input. Some experiments have found positive correlations between the strength of preparatory target-selective activations in visual cortex and target detection performance (Giesbrecht et al., 2006; Stokes et al., 2009; Peelen & Kastner, 2011). However, other studies have failed to observe any systematic relationships between anticipatory baseline shifts and subsequent target-selective modulations of visual processing or behavioural target selection efficiency (e.g., Fannon, Saron, & Mangun, 2007; McMains, Fehd, Emmanouil, & Kastner, 2007). The difficulty in finding strong causal links between preparatory activity in visual-perceptual brain areas and the selective attentional processing of targets versus distractors raises doubts about the idea that these baseline shifts are the direct neural counterpart of search templates. Even if such baseline shifts are linked to the maintenance of target-related information in working memory, they may reflect a type of working memory that is not suited to the functions of attentional templates.

To understand why this might be the case, it is important to consider how information about visual objects is represented in working memory. Such representations can be either position-dependent or position-invariant (spatially global). In position-dependent representations, the spatial layout of visual information that is encountered during encoding is retained. In contrast, position-invariant representations contain no explicit information about particular object locations in the visual field. This distinction is important when considering the role of working memory representations as attentional templates in visual search tasks where the position of a target object among distractors in the visual field is uncertain. If the function of attentional templates is to affect the subsequent allocation of attention in a goal-selective fashion, these templates need to operate in a spatially global fashion across all possible target locations in the visual field. In other words, attentional templates should be position-invariant. If working memory representations that are reflected by preparatory baseline shifts in posterior visual areas were strongly position-dependent, this would be inconsistent with their role as attentional templates during search for known targets at unknown locations.

There are two reasons why working memory representations in visual cortex should be position-dependent. On the one hand, this prediction follows directly from the sensory recruitment account of working memory which assumes that memorized objects are maintained in posterior visual regions that are also responsible for the perceptual analysis of

incoming visual signals. In visual cortex, information is represented in a position-dependent fashion in two-dimensional cortical maps (e.g., Franconeri, Alvarez, & Cavanagh, 2013). In these maps, the spatial coordinates of visual features and objects are defined relative to their position on the retina (retinotopic representation) or in the external world (spatiotopic representation), and even higher-level visual areas retain strong retinotopic biases (e.g., Desimone & Gross, 1979; Op De Beeck & Vogels, 2000; Kravitz, Kriegeskorte, & Baker, 2010; see Kravitz et al., 2013, for a review). A second reason for the position-dependence of working memory representations that will be discussed in more detail in section 5 is that memory maintenance is mediated by focal spatial attention, which necessarily operates on space-based representations of visual objects.

Perhaps the most direct evidence for position-dependent memory representations in visual cortex comes from ERP studies which showed that neural activity during the delay period of working memory tasks is elicited at posterior electrodes contralateral to the side where the to-be-remembered objects appeared during encoding (contralateral delay activity/CDA; see Vogel & Machizawa, 2004). In fact, there is a direct somatosensory analogue of the visual CDA (tactile CDA component; see Katus, Grubert, & Eimer, 2014) that shows a distinct modality-specific topography over lateral somatosensory cortex. Such ERP results demonstrate that the spatial layout of sensory information is retained when this information is stored and maintained in working memory (see also Hornak, Duncan, & Gaffan, 2002; and Gratton, Corballis, & Jain, 1997, for additional behavioural evidence for the position-dependence of visual memory). At the neural level, such position-dependent visual working memory representations should be implemented by object-selective sustained activity modulations at particular locations within visual cortical maps. If target-selective baseline shifts of visual activity that have been observed during the preparation for visual search show this kind of location-specificity, they would not be able to modulate subsequent visual processing in a spatially global fashion. This may be the primary reason why it has proved difficult to demonstrate causal links between preparatory activity modulations in visual cortex and subsequent stages of attentional processing.

Instead of being position-dependent, attentional templates in visual search need to represent search targets in a spatially global position-independent fashion. It is possible that such spatially global search templates may not be found at all in visual-perceptual areas, but only in higher-level attentional control regions such as prefrontal cortex, where visual

information is represented in a largely position-independent fashion. In fact, patterns of neural activity in prefrontal cortex that are sensitive to current search targets have indeed been observed during the preparation for attentional tasks (Peelen & Kastner, 2011; Warden & Miller, 2010; Stokes et al., 2013). However, it might be premature to completely rule out visual cortex as a possible additional locus for position-independent preparatory attentional templates. In a study by Ester, Serences, & Awh (2009), participants memorized the orientation of a grating in the left or right visual field during a delay period before matching it to a test grating. Pattern analyses of fMRI data obtained during the delay period found that activity in contralateral primary visual cortex at locations that matched the memorized grating was sensitive to its orientation, as would be expected if working memory was position-dependent. Critically, Ester et al. (2009) found that corresponding areas of ipsilateral primary visual cortex were equally sensitive to the memorized orientation. This indicates that orientation information was maintained in a spatially global fashion, and suggests that position-independent working memory representations may also exist in visual cortex. If this is the case, such representations could act as spatially global attentional templates in visual search.

The hypothesis that visual working memory representations can either be position-dependent or position-invariant, and that only position-invariant representations can act as attentional templates during visual search, might also explain another apparent dissociation between search templates and other types of visual working memory representations. Visual working memory has a capacity of approximately 3-4 objects (Luck & Vogel, 1997; Cowan, 2001). If search templates are stored in working memory, it should in principle be possible to simultaneously activate multiple attentional templates for different possible target features or objects, up to the point where memory capacity is exceeded. In fact, search for multiple targets is much less efficient than search for one particular object or feature. Houtkamp & Roelfsema (2009) demonstrated that the detection of targets in a rapid serial visual presentation (RSVP) stream is impaired when observers search for one of two possible target objects relative to a task where they searched for a single object. Modelling of these behavioural results suggested that exactly one attentional template can be active at any given time (see also Meneer, Cave, & Donnelly, 2009, and Stroud et al., 2011, for additional behavioural evidence, and Grubert & Eimer, 2013, for ERP evidence that attentional target selectivity is less effective during multiple-colour as compared to single-

colour visual search). If attentional templates are working memory representations, and if working memory can hold several objects simultaneously, why should only a single search template be active at a time? One possibility is that attentional templates and other working memory representations are structurally equivalent, except that the current search template is more strongly activated, and is therefore able to bias focal attentional processing towards template-matching objects (Olivers et al., 2011). Other working memory representations can be simultaneously maintained, but are temporarily inhibited and therefore unable to affect the allocation of attention (see also Olivers & Eimer, 2011). Alternatively, the difference between an attentional template and other visual working memory items might reflect a more fundamental qualitative difference between position-invariant and position-dependent representations. In this context, the impaired efficiency of multi-target as compared to single-target search would suggest that it is difficult to maintain more than a single spatially global attentional template at any given time.

In summary, the research discussed in this section has focused on the nature of preparatory search templates and their neural basis. It is generally agreed that attentional templates are representations of target objects or features that are maintained in working memory during the preparation for visual search. However, identifying the neural correlates of such templates and demonstrating causal links between preparatory neural activity and subsequent attentional effects has proved to be difficult. According to the sensory recruitment model of visual working memory, search templates should reside in visual-perceptual cortical areas, but this is complicated by the fact that information in these visual areas is represented in position-dependent cortical maps. Because search templates have to operate in a spatially global fashion, they should represent search targets irrespective of their particular location in visual space. Such position-independent representations of target objects exist in prefrontal cortex, and possibly also in visual areas, and these representations might be the neural counterpart of preparatory attentional templates in visual search.

### 3. Attentional guidance and feature-based attention

While attentional templates are set up in preparation for an upcoming search task, the search process itself starts once a visual search display has been presented. When the location of search targets is unknown, focal attention cannot be allocated to a particular region of visual space during the preceding preparatory phase. The selection of possible target objects therefore needs to be based on the visual information that is available in the search display itself. According to models of visual search (e.g., Treisman & Sato, 1990; Wolfe, 1994, 2007), this information is accumulated at early stages of visual processing. For example, the Guided Search model (Wolfe, 1994, 2007) assumes that representations of currently task-relevant object features are positively weighted during the initial parallel processing of visual input, thereby increasing the probability that these features will attract focal attention. Similar ideas were proposed in the biased competition account of selective attention (Desimone & Duncan, 1995; Duncan, 2006). According to this account, multiple simultaneously active object representations compete for neural processing resources and the control of behaviour, and this competition is biased in favour of currently task-relevant objects that are specified by attentional templates (i.e., preparatory baseline shifts of neural activity). These attentional biases already operate at early stages of visual processing, although it may take considerably longer before the competition for attentional selection is resolved.

If the function of such goal-selective biases is to guide the allocation of attention during search for known targets at unknown locations during the early parallel processing of visual input, these biases need to operate in a spatially global fashion across the visual field. There is indeed considerable evidence for the existence of such spatially global attentional biases from research on feature-based attention. Results from single-unit studies in monkeys have shown that when a specific visual feature is currently task-relevant, the neural processing of this feature is enhanced at the expense of the processing of other features in the same dimension. Crucially, these feature-based attentional modulations of neural activity appear to be elicited in a spatially global fashion across the entire visual field. In a study by Martinez-Trujillo & Treue (2004), two sets of dots that both moved in the same direction were presented in the left and right visual field, and monkeys were trained to detect small changes in the speed and direction of one these sets of dots. The activity of

movement-selective neurons in the middle temporal cortical area with receptive fields that covered the stimuli in the currently unattended visual field was strongly modulated by the direction of movement that the monkey was attending on the other side. Neurons that preferred the currently task-relevant movement direction showed enhanced activation, while the activity of neurons that preferred the opposite direction of movement was suppressed. In other words, these attention-dependent activity modulations of movement-sensitive neurons were triggered in response to stimuli in the task-irrelevant unattended visual field. Further evidence that feature-based attention operates in a spatially global fashion across the visual field was provided by Bichot, Rossi, & Desimone (2005) in a study where monkeys searched for colour-defined or shape-defined target objects. Neurons in visual area V4 that were selective for the currently task-relevant feature increased their activity when a target object was present in their receptive field even when the monkey fixated a different object, and then shifted eye gaze to another location. This demonstrates that goal-dependent attentional modulations are elicited outside the current focus of attention, thus providing further evidence that this type of feature-based visual processing bias operates in a spatially global fashion.

Additional evidence for the existence of spatially global feature-based attentional modulations was provided by fMRI and ERP studies in humans (Saenz, Buracas, & Boynton, 2002; Serences & Boynton, 2007; Zhang & Luck, 2009). In these studies, observers attended to a specific task-relevant feature in one visual field, and objects in the other unattended hemifield triggered enhanced visual responses when they matched the feature that was currently attended on the opposite side. Although such spatially global effects of feature-based attention have mostly been observed for simple target features such as colour, shape, or movement direction, they may also be present during search for category-defined targets. This has been demonstrated by Peelen, Fei-Fei, & Kastner (2009), who asked participants to report the presence of people or cars in briefly presented images of real-world visual scenes at a particular location, and to ignore images that were simultaneously presented at other irrelevant locations. MVPA-based analyses of fMRI data revealed neural response patterns in object-selective visual cortex that were sensitive to the presence of the currently task-relevant stimulus category in a particular image, even when this image appeared at a to-be-ignored location. This suggests that spatially global modulations of

visual processing in favour of possible target objects are triggered not only in simple feature-based attentional selection tasks, but also during category-based search.

If feature-based attention operates in a spatially global fashion, its utility for the control of attentional selectivity in visual search is obvious. Feature-based attention can bias perceptual processing in favour of candidate target objects, irrespective of the location that these objects occupy in the visual field, and can thus direct spatial attention to objects that match one or more currently task-relevant features. In fact, this type of spatially global feature-based attentional control may represent the direct neural equivalent of the guidance component postulated in the Guided Search model (Wolfe, 1994, 2007). In this model, the parallel processing of visual information across the entire visual field is selectively weighted in favour of features that match current search goals, resulting in the allocation of focal attention towards likely target objects (see also Bundesen et al., 2005, for similar ideas). It is interesting to note that some target-defining visual attributes are much more effective than others in facilitating attentional object selection during visual search (see Wolfe & Horowitz, 2004, for a review). If the selection of search targets is guided by spatially global feature-based attentional modulations, such differences could reflect differences in the availability of goal-selective biases during the parallel analysis of visual information. Due to the modular architecture of the visual system, spatially global feature-specific attentional biases might be relatively easy to implement for simple target features such as colour, motion, or orientation, but not for more complex target-defining attributes such as line intersections or three-dimensional volume. However, as shown by the presence of spatially global processing modulations during category-based search (Peelen et al., 2009), this does not necessarily imply that feature-based attentional control is restricted to elementary visual features. The possibility that feature-based attentional mechanisms are more readily available for some target-defining attributes than others, and that this determines the effectiveness of top-down attentional control in visual search obviously requires further empirical support.

What are the links between the spatially global feature-based attention effects that are observed during early stages of visual processing and the target-selective baseline shifts that are elicited during the preparation for a visual search task? In other words, how do the preparatory attentional templates that were discussed in the previous section affect attentional guidance processes? If attentional templates represent search targets in a

spatially global fashion, it is plausible to assume that these templates are directly responsible for the emergence of spatially global feature-based attentional modulations. Such causal links between preparation and guidance may be either direct or more indirect. If attentional templates are position-independent representations of search targets in visual cortex (as suggested by recent evidence for spatially global working memory representations in visual cortex; Ester et al., 2009), the activation of such representations during the preparation phase of visual search may be simply maintained once search displays are presented. In this case, attentional templates in visual working memory that are set up during the preparation for search and feature-based attentional modulations that are observed once visual input has been received would be essentially two sides of the same coin (see also Desimone & Duncan, 1995, for similar suggestions). The observation that feature-based attention effects can spread to currently empty regions of visual space (Serences & Boynton, 2007) suggests that preparatory feature-selective biases that are already active prior to the arrival of visual input can persist after a search display is encountered. The sustained presence of preparatory baseline shifts can modulate the rapid feedforward processing of visual input. This will result in an enhancement of visual activation at all locations of template-matching objects in the visual field, as reflected by spatially global feature-based attention effects described earlier. For example, the spatially global modulations of category-selective responses observed by Peelen et al. (2009) in response to search displays containing multiple images may directly reflect the persistence of category-specific search templates that were set up during the preceding preparation phase (e.g., Peelen & Kastner, 2011). Alternatively, if spatially global attentional templates reside not in visual cortex, but instead in more anterior areas such as prefrontal cortex where search targets are represented in a position-independent fashion, the links between preparation and guidance may be more indirect. In this case, feature-based attention effects may be initiated and controlled by top-down signals from prefrontal to visual areas (e.g., Maunsell & Treue, 2006). These two alternative scenarios are not necessarily mutually exclusive, as goal-directed visual selection processes will generally involve bidirectional recurrent interactions between visual-perceptual brain regions and more anterior attentional control areas (e.g., Bundesen et al., 2005).

Overall, the research discussed in this section strongly suggests that when observers search for a particular target at an unknown location, even early parallel stages of visual



information processing are already modulated by search goals that have been activated during the preparation for search. Because such feature-based attention effects operate in a spatially global fashion, they can highlight the presence of potential target objects anywhere in the visual field, and provide guidance signals for the subsequent allocation of focal attention to particular objects. More generally, the presence of goal-sensitive attentional modulations during the rapid parallel analysis of visual input is interesting because it questions the assumptions of traditional two-stage models of visual perception and attention. Such models distinguish between an initial pre-attentive processing stage that is entirely driven in a bottom-up fashion by the physical properties of the visual input and a second attentive stage where visual processing is affected by current selection intentions. This two-stage scenario was first proposed by Broadbent in his filter theory of selective attention (1958), and has been a key feature of virtually all theoretical accounts of visual perception and attention ever since. For example, Theeuwes (2010) proposed a model where the initial stage of attentional selection is determined exclusively by bottom-up salience signals generated during pre-attentive vision, and top-down influences only emerge at later stages of attentional processing. The possibility that feature-based attention already affects the early parallel feedforward analysis of visual input in a goal-dependent fashion casts serious doubts on the existence of a distinct pre-attentive stage of visual processing that operates in a genuinely stimulus-driven non-selective fashion, and thus on the validity of the fundamental distinction between pre-attentive and attentive vision.

#### **4. Object selection and focal spatial attention**

Feature-based attention operates in a spatially global fashion during the initial parallel processing of visual input, and highlights the presence of potentially task-relevant objects across the visual field. This shows that even early stages of visual information processing are already selective in the sense of reflecting specific search goals. However, most models of visual search (e.g., Treisman & Gelade, 1980; Wolfe, 1994, 2007) assume the existence of a separate and distinct attentional object selection process. Although the concept of selection plays a central role in theories of attention, its precise meaning is rarely made explicit. In traditional two-stage models of visual processing, 'selection' marks the

transition from pre-attentive to attentive vision. Such models explain the need for object selection by reference to the limited capacity of attentional processing, and characterize the selection process as a gatekeeping mechanism that regulates access to this limited-capacity system (e.g., Broadbent, 1958). This account of selection and its link to generic cognitive capacity limitations has been criticised by other theorists (e.g., Allport, 1993). More recently, the limited capacity of visual processing and the resulting need for attentional selectivity have been described more specifically as a direct result of the space-based topographical organisation of the visual system. Because visual objects are represented in two-dimensional cortical maps, multiple objects can share the same neuronal receptive fields. These objects will compete for representational space (“cortical real estate”; Franconeri et al., 2013), that is, for the control of neural responses at particular locations within the visual maps. In this scenario, capacity is limited in the sense that only a small number of objects can be neurally represented at any given time. ‘Selection’ refers to the outcome of a competitive process, where a particular object has succeeded in driving neural activity at a particular location of visual space (Desimone & Duncan, 1995; Duncan, 2006). In this context, attentional object selection is space-based, and is defined as the emergence of spatially focal processing biases in favour of particular objects, at the expense of other simultaneously present competing objects.

An electrophysiological signature of this type of object selection was described by Chelazzi et al. (1998). In this study, monkeys had to make a saccade to a specific target object that was previously specified by a picture cue and was accompanied by a nontarget distractor object on the same side. As described in section 2, the picture cues elicited sustained baseline shifts of visual activity during the cue-target interval that were interpreted as neural correlates of preparatory attentional templates. Neural responses to the subsequent target/nontarget displays were recorded from inferotemporal cortex, for neurons that preferred one of these two objects in these displays. Response rate was initially high, regardless of whether this preferred object was the saccade target or the distractor on a given trial, due to the presence of the preferred object in the receptive field. However, from about 180 ms after search display onset, neural responses were determined by search goals. When the preferred object was the saccade target, response rate remained high. In contrast, neural activity decreased strongly on trials where the same object served as distractor. In line with previous observations that spatial attention determines the

response rate of visual neurons when task-relevant and irrelevant stimuli are simultaneously present in their receptive fields (e.g., Moran & Desimone, 1985), these results of Chelazzi et al. (1998) suggest that a spatially selective bias in favour of the target object emerged within less than 200 ms after stimulus onset. This spatial bias can be interpreted as the neural correlate of attentional object selection. Very similar electrophysiological effects have been found in many ERP studies of visual search in humans. When a candidate target object in the left or right visual field is presented together with distractors, this object triggers an enhanced negativity at contralateral occipito-temporal electrodes (N2pc component; Luck & Hillyard, 1994; Eimer, 1996; Girelli & Luck, 1997). Similar to the target-selective modulations of neural responses described by Chelazzi et al. (1998), the N2pc component typically emerges around 180 ms after stimulus onset, and is primarily generated in extrastriate ventral visual cortex (e.g., Hopf et al. 2000). It is therefore likely that both measures are linked to the same underlying neural process. Both reflect the spatially selective enhancement of responses to potential target objects versus distractors in visual areas, and both are neural markers of attentional object selection processes in visual search.

Most ERP studies that have used the N2pc component as an electrophysiological marker of attentional object selection in visual search have investigated situations where search targets were defined by simple visual features such as particular colours or shapes. In such tasks, the N2pc emerges within less than 200 ms after search display onset, demonstrating that spatially specific attentional object selection processes are elicited during relatively early stages of visual processing. Even though one might assume that attentional selection operates much more slowly during search for category-defined targets, a series of recent N2pc experiments from our lab suggest that this is not the case. In one study, participants searched for targets that were defined with respect to their alphanumeric category (e.g., any letter among digit distractor objects; Nako, Wu, & Eimer, 2014). N2pc components to target items were elicited at around 180 ms post-stimulus, which is very similar to the N2pc onset latencies typically observed in search tasks where targets are defined by simple visual features. In another study, targets were line drawings of real-world objects that were defined in terms of their category membership (e.g., kitchen objects among items of clothing; Nako, Wu, Smith, & Eimer, 2014). Here, the N2pc to target objects emerged slightly later, at around 240 ms post-stimulus. The observation that category-based attentional selection processes are triggered within less than 250 after a

search display has been presented shows that category search mechanisms can operate remarkably rapidly. This is consistent with the results from fMRI studies that investigated preparation and guidance processes during visual search for category-defined real-world target objects. Category-selective modulations of activity in visual cortex were already observed when participants prepared for a particular search episode (Peelen & Kastner, 2011), and similar effects were elicited in a spatially global fashion during the subsequent processing of visual input (Peelen et al., 2009). These findings suggest that attentional preparation and subsequent attentional guidance processes can be selectively set for specific target categories, which may explain the rapid emergence of category-based attentional selection processes that was observed in our recent N2pc studies.

What is the relationship between the spatially global attentional guidance mechanisms discussed in the previous section, and the attentional selection processes discussed here? The attentional selection of a particular object during visual search (i.e., the emergence of a spatially selective bias in favour of this object) is assumed to be based on information about the locations of candidate target objects that is accumulated during the preceding guidance phase (Wolfe, 1994, 2007). Different architectures have been suggested to describe this interplay between guidance and selection. In hierarchical models, potential target locations are represented via priority or salience maps (e.g., Itti & Koch, 2001; Fecteau & Munoz, 2006). Information about the presence of potential targets is generated in parallel for different feature dimensions, and converges on a shared priority map where it is integrated (e.g., Wolfe, 2007). The priority map is located in dedicated attentional control regions that are anatomically and functionally distinct from the visual areas where spatially selective biases towards target objects emerge. The frontal eye fields (FEFs), posterior parietal cortex, or the thalamus have all been considered as the potential neural locus of such maps (e.g., Schall, 2004; Gottlieb, Kusunoki, & Goldberg, 1998; Bundesen et al., 2005). Specific locations in a priority map are linked to spatially corresponding locations in visual cortex, so that information about likely target locations within this map can trigger spatially selective modulations of visual processing. This hypothesis is supported by the observation that electrical stimulation of FEF (one of the potential attentional control areas) modulates the activity of spatially corresponding regions in visual area V4 (Moore & Armstrong, 2003). In terms of the distinction between guidance and selection, the creation of a specific priority map can be described as the result of a spatially global guidance mechanism, and the

subsequent selective modulation of visual activity at particular locations as the resulting attentional selection process.

Not all accounts of attentional control during visual search postulate the existence of dedicated priority maps. According to non-hierarchical distributed models, goal-sensitive biases in favour of particular objects or features can be generated at different levels of the visual processing hierarchy. These biases are then propagated both to higher and lower levels where spatially selective competitive advantages for specific objects emerge ("integrated competition"; Duncan, Humphreys, & Ward, 1997). In these models, the transition from guidance to selection during visual search is a continuous process where the competition between multiple objects is gradually resolved in favour of those objects with currently task-relevant properties.

A controversial issue in past and present debates about the mechanisms of attentional object selection in visual search concerns the serial versus parallel nature of such selection processes. Does attentional selection operate sequentially for one object at a time, or in parallel, so that several objects can be selected simultaneously and independently? If selection is implemented at the neural level as the emergence of spatially specific activity modulations of object representations at particular locations in visual maps, the issue of parallel versus serial selection refers to the question whether such modulations will eventually be restricted to one specific location or can be maintained simultaneously at multiple locations in the visual field. In traditional two stage-models of visual perception and attention, the transition from pre-attentive to attentive vision that is controlled by selection mechanisms coincides with the transition from parallel to serial processing. This assumption is retained in current models of visual search which describe attentional object selection as a serial process. According to Feature Integration Theory, focal attention is directed sequentially to individual objects, which implies that the attentional selection of a new object is preceded by a de-allocation of attention from its previous location (e.g., Treisman & Gelade, 1980). Along similar lines, the Guided Search model (Wolfe, 1994, 2007) describes the selection of individual objects during visual search as a serial attentional bottleneck. Perhaps the most compelling reason for assuming that object selection operates in a serial fashion is that the allocation of spatial attention is closely linked to eye movement control (e.g., Moore, Armstrong, & Fallah, 2003; Thompson & Bichot, 2005). Saccadic eye movements are executed sequentially, and are preceded by the allocation of spatial

attention to the location of the next saccade target (e.g., Deubel & Schneider, 1996). Such attention shifts that precede eye movements are necessarily serial. However, spatial attention can be allocated in the absence of overt gaze shifts, and visual search does not depend on eye movements (e.g., Zelinsky & Sheinberg, 1997). For these reasons, the serial nature of oculomotor control processes is not sufficient to conclude that attentional object selection always operates in a strictly serial fashion. In fact, several theories of attention (e.g., Desimone & Duncan, 1995; Bundesen et al., 2005) postulate that object selection processes can operate in parallel at multiple locations in the visual field. Along similar lines, the ability to simultaneously track multiple moving objects in the visual field has been explained by assuming that focal spatial attention is allocated independently and in parallel to these objects (Cavanagh & Alvarez, 2005).

As an electrophysiological marker of attentional object selection, the N2pc component can track the time course of attentional selection processes in visual search on a millisecond-by-millisecond basis, and can therefore provide insights into the parallel versus serial nature of these processes. Indirect N2pc evidence for parallel selection comes from the observation that N2pc amplitudes are sensitive to the number of task-relevant objects in a display. When observers have to report how many colour-defined target objects are present among distractors on one side of a search display, N2pc amplitudes increase with the number of targets (Mazza & Caramazza, 2011; see also Drew & Vogel, 2008, for similar observations). This N2pc amplitude increase has been interpreted in terms of object individuation processes, which operate in parallel when multiple objects have to be simultaneously distinguished from distractor objects. If ‘object individuation’ is the same process as ‘object selection’, this would imply that the increase of N2pc amplitudes with the number of targets in a search display reflects attentional selection processes that operate simultaneously and in parallel for each target object.

More direct N2pc evidence for the existence of such parallel object selection processes was obtained in a recent study from our lab (Eimer & Grubert, 2014a) where the selection of one target was measured independently of the selection of another target. In this study, two displays were presented in rapid succession, with stimulus onset asynchronies (SOAs) of either 100 ms or 10 ms (as shown in Figure 2, top panel). Both displays contained one target object that was defined by one specific colour, and was accompanied by a distractor object in a different colour on the opposite side. Participants

had to identify the two target objects that were successively presented in display 1 and display 2, and to report whether these targets belonged to the same alphanumerical category (two letters, two digits) or not (one letter and one digit). The target/nontarget pair in one display always appeared on the horizontal meridian (to the left and right of fixation), while the stimulus pair in the other display was presented on the vertical meridian. This procedure was used to track the attentional selection of one of the two target objects independently of the selection of the other target. Because the N2pc is always elicited contralateral to the side of a target object in the left or right visual field, no N2pc is triggered by objects that appear on the vertical meridian. In our study, N2pc components therefore always reflected the attentional selection of the horizontal target. When the two displays were separated by a 100 ms SOA, the N2pc to horizontal targets in display 1 preceded the N2pc to horizontal targets in display 2 by almost exactly 100 ms. The N2pc to targets in the first display was followed by a second negative peak that overlapped with the N2pc to the targets in display 2 (Figure 2, SOA 100, right panel). This second peak reflects the initial phase of the sustained contralateral negativity that is associated with the encoding of task-relevant stimuli into working memory (see section 5). Critically, when the two displays appeared within 10 ms of each other, the N2pc to horizontal targets in display 2 emerged 10 ms later than the N2pc to target in display 1 (Figure 2, SOA10). These two N2pc components were equal in size and overlapped in time, demonstrating that spatial attention was allocated rapidly and in parallel to both target objects, with each selection process following its own independent time course (see also Khayat, Spekreijse, & Roelfsema, 2006, for corresponding evidence for temporally overlapping attentional selection processes from monkey neurophysiology). Similar results were obtained in another set of studies that employed the same procedures, except that the two target objects were now defined by two different colours (Grubert & Eimer, 2015), so that their selection could no longer be controlled by attentional preparation and guidance processes that are set for a single target colour. In spite of this fact, the temporal pattern of N2pc components was similar to the pattern observed in our initial experiment (Eimer & Grubert, 2014a). These N2pc results are difficult to reconcile with a strictly serial selection account, which implies that the selection of a new object can only commence once attention is withdrawn from its previous locus. They suggest instead that multiple attentional selection processes can operate independently and in parallel.

Additional evidence for this conclusion comes from a visual search study from our lab (Eimer & Grubert 2014b) where participants searched for a target that was defined by a specific conjunction of colour and shape (e.g., a blue circle). This target was presented together with two task-irrelevant distractors and with an additional nontarget object that matched one of the two target-defining features (e.g., a blue square). On different trials, the target object was presented on the horizontal meridian and the partially matching nontarget object on the vertical meridian, or vice versa, so that N2pc components could be measured independently to both types of objects. According to the Guided Search model (Wolfe, 2007), attention should always be allocated to the target object, because this object has both task-relevant features and will therefore trigger the strongest activation on the priority map. If this is correct, N2pc components should be elicited only by targets, but not by a partially matching nontarget object in the same display. In fact, reliable N2pc components were observed not only for targets, but also for partially matching nontargets, even though the target object was simultaneously present. This shows that attention was allocated in parallel and independently to all features in the display that matched the current target attributes (see also Andersen, Hillyard, & Müller, 2008, for corresponding evidence for the parallel selection of target features from steady-state visual evoked potentials).

Overall, these N2pc results show that attentional object selection processes can be elicited in parallel for multiple objects with target-matching features. Such observations are difficult to reconcile with the widely held view that selection operates serially in visual search (e.g., Treisman & Gelade, 1980; Wolfe, 2007). But does the N2pc component exclusively reflect processes that operate during the attentional selection phase? It is possible that the N2pc might also be sensitive to processes that take place during the earlier spatially global attentional guidance stage. In fact, the pattern of N2pc results observed in our recent studies where different objects with target-matching features were presented simultaneously or in rapid succession (Eimer & Grubert, 2014a, 2014b; Grubert & Eimer, 2015) appears to be similar to the spatially global modulations of visual processing produced by feature-based attention that were described in the previous section as the neural correlate of attentional guidance. Although such feature-based attention effects typically emerge earlier than the N2pc component (e.g., Hopf et al., 2004; Zhang & Luck, 2009), the question remains whether the N2pc results discussed earlier reflect spatially global feature-



based attentional guidance or parallel space-based attentional selection processes. This question assumes that there is a strict separation between spatially global guidance processes (feature-based attention) and spatially specific selection mechanisms (spatial attention), and that this dichotomy describes functionally and temporally discrete stages of attentional processing during visual search. In fact, the distinction between guidance and selection is a heuristically useful way of conceptualising different aspects of attentional selectivity, but does not reflect the essentially continuous nature of attentional mechanisms at the neural level. Because attentional selectivity develops gradually during the processing of visual input, the transition from spatially global attentional guidance to spatially focal attentional object selection is a continuous process, where early feature-based spatially global biases of visual processing gradually develop into object-based spatially selective processing modulations.

In summary, attentional object selection was described in this section as the emergence of spatially specific visual processing biases in favour of particular objects that result from the competition for representational space in visual cortical maps. These biases are typically elicited within less than 200 ms after stimulus onset when search targets are defined by simple visual features, and can also emerge rapidly during category-based visual search. They are the result of information about the presence of possible target objects that is accumulated by attentional guidance processes. Spatially selective modulations of visual processing can be triggered simultaneously and independently at different locations in the visual field, which implies that attentional selection processes can operate in parallel for different objects.

## **5. Object recognition and working memory**

The neural basis of attentional object selection was described as the emergence of a spatially specific bias of visual processing within visual cortical maps in favour of a particular task-relevant object. However, the presence of such object-selective attentional modulations does not imply that selected objects are instantly recognized. During the attentional tracking of multiple visual objects, access to the features and identity of these objects is remarkably poor (Horowitz et al., 2007), demonstrating that the allocation of focal

attention to specific objects is not sufficient for their recognition. In visual search, the selection and identification of target objects are separable mechanisms (e.g., Ghorashi, Enns, Klein, & Di Lollo, 2010). Many models of visual attention and visual search make an explicit distinction between object selection and object recognition. Object selection is described as a stage where particular objects are individuated via the allocation of focal attention. Object recognition is assumed to take place at a subsequent stage where the features of these objects are integrated and their identity becomes accessible (e.g., Wolfe, 2007; Huang & Pashler, 2007; Xu & Chun, 2009). In line with these suggestions, recent studies from our lab have also found ERP evidence for the transition between an early phase of attentional selectivity where attention is rapidly allocated to candidate target objects and a later phase where information about the attributes of selected objects is integrated across feature dimensions (Kiss, Grubert, & Eimer, 2013; Eimer & Grubert, 2014b).

Selection and recognition are sensitive to different experimental factors: The efficiency of target selection is determined by the number of competing nontarget objects in a search display and their similarity to the target (e.g., Duncan & Humphreys, 1989), whereas recognition processes are primarily affected by target complexity (e.g., Xu & Chun, 2009; Franconeri et al., 2013). When two target objects are presented successively and without competing distractors in the same display, so that the demands on spatial selectivity are minimal, identification of the second target is often strongly impaired ("attentional blink"; e.g., Duncan, 1980; Raymond, Shapiro, & Arnell, 1992; Duncan, Ward, & Shapiro, 1994; see Wyble, Bowman, & Nieuwenstein, 2009, for an account of the attentional blink in terms of competitive mechanisms in working memory). This suggests that attentional capacity limitations can arise specifically at object recognition stages that follow the spatial selection of target objects.

How can the transition from object selection to object recognition be described at the cognitive and neural levels? The recognition of a particular object is based on the activation of a working memory representation of this object (e.g., Bundesen et al., 2005; Chun & Johnson, 2011). Such working memory representations are actively maintained by a sustained focus of spatial attention (e.g., Oberauer, 2002; Awh, Vogel, & Oh, 2006). Object selection was previously defined as the allocation of spatial attention to particular objects, as reflected by a spatially selective modulation of neural responses to these objects in visual cortical maps. The encoding of a selected object into working memory during the

recognition phase can thus be described as the active maintenance of an attentional focus that was established during object selection, or, in other words, as spatial attention that is sustained internally over time (e.g., Chun, 2011).

This critical role of spatial attention for visual working memory was shown in behavioural and ERP studies that found spatially selective visual processing enhancements for locations that were currently maintained in memory (Awh, Jonides & Reuter-Lorenz, 1998; Awh, Anillo-Vento, & Hillyard, 2000). According to the sensory recruitment model of working memory (e.g., Postle, 2006), memorized visual objects are stored in visual areas that are also activated during the perceptual processing of visual input. The attention-based maintenance of visual objects should therefore take place within cortical maps in visual areas, and should operate via spatially selective activity enhancements for visual object representations at particular locations within these maps, because spatial attention necessarily operates in a space-based fashion. This essential involvement of spatial attention in the maintenance of working memory representations is one of the reasons why visual working memory representations are strongly position-dependent (see section 2).

If object selection and working memory are both mediated by spatial attention, this should be reflected by functional links between the N2pc component (which marks attentional object selection; see section 4) and the subsequent CDA component (which is elicited during working memory maintenance). As mentioned in section 2, CDA components are elicited during the delay period of working memory tasks at posterior electrodes contralateral to the side where memorized objects were presented during encoding (Vogel & Machizawa, 2004). In a study where observers had to memorize between one and six display objects for subsequent recall (Anderson, Vogel & Awh, 2011), CDA amplitudes increased with memory set size, up to the point where individual working memory capacity was exceeded. Importantly, N2pc components that were triggered during the initial attentional selection of the to-be-remembered objects showed exactly the same sensitivity to memory set size. Such parallel effects of memory set size on N2pc and CDA amplitudes should indeed be observed if separate independent foci of spatial attention are established during the attentional selection of memorized objects (N2pc component), and are then sustained over time during the maintenance of these objects in working memory (CDA component).

The hypothesis that attentional object selection and working memory maintenance are both based on spatially selective modulations of visual processing at one or several locations in the visual field yields another interesting prediction: Individual differences in the ability to maintain multiple objects in visual working memory should be linked to individual differences in the ability to allocate spatial attention simultaneously to multiple objects during visual search. The existence of such a link between memory capacity and search performance was demonstrated by Anderson, Vogel, & Awh (2013). Individuals with high working memory capacity performed more efficiently than low-capacity participants in a difficult search task where targets and distractors were very similar, so that each item had to be focally attended in order to be recognized as target or nontarget. These observations suggest that object selection during visual search tasks and working memory capacity are determined by a common underlying factor - the ability to select and maintain multiple representations of individuated objects in a spatially selective fashion.

The critical role of focal spatial attention during the selection and subsequent active maintenance of objects in visual working memory is illustrated by a recent ERP study from our lab (Towler & Eimer, in press) that investigated visual face memory. Memory displays that contained two different faces in the left and right visual field were followed after a brief delay period by test displays with a single face at fixation (see Figure 3, left panel). Participants' task was to encode and maintain both faces in the memory display, and to report whether the test face matched one of these two faces or was a different face. Performance was surprisingly poor in this task, and suggested that only one of the two faces in the memory displays was encoded into working memory on most trials. Which of these two faces was maintained was determined by spatial attention, and this was revealed by the ERPs recorded in response to the memory displays, prior to the arrival of the subsequent test displays. Figure 3 (right panel) shows these ERPs for trials where one of the two memory display faces was later repeated in the test display. On trials where participants detected this face repetition correctly and rapidly, N2pc and CDA components were found contralateral to the face in the memory display that was then repeated. On these trials, focal attention was allocated to the 'correct' face (i.e., the face that would reappear as the test face), and this spatial focus was then maintained during the delay period, resulting in the rapid detection of an identity match between the memorized face and the test face. On trials where participants failed to detect a face repetition, N2pc and CDA components were

instead elicited contralateral to the face in the memory display that was not repeated. On these trials, focal attention was evidently allocated to the 'wrong' (i.e., non-repeated) face, and this face was then retained in working memory, at the expense of the face that would reappear as the test face. These ERP results show that the success or failure of maintaining a visual representation of an individual face in working memory is determined by the allocation of spatial attention, and that attention can only be allocated to one face at a time.

If object selection is implemented through the spatially selective modulation of neural responses to particular objects in cortical maps, and object recognition is based on the active maintenance of this spatial bias over time, the question arises whether these two successive stages can be dissociated. The maintenance of a particular object in working memory during its recognition should always depend on the prior attentional selection of this object. But are there situations where objects are selected without subsequently being encoded in working memory? Electrophysiological evidence for selection without memory-based recognition was found in a visual search study from our lab (Mazza, Turatto, Umiltà, & Eimer, 2007) where observers had to select a colour singleton target among uniform distractor objects in two different task conditions. In a localisation task, they simply had to report whether the target appeared in the left or right visual field. In a discrimination task, the specific shape of the colour-defined target object had to be identified. Identical N2pc components were triggered in both tasks, demonstrating that the initial attentional selection of targets was unaffected by the difference in task demands. In contrast, the subsequent sustained contralateral negativity that marks the activation of a target representation in working memory mediated by focal spatial attention was only elicited when participants had to discriminate the target shape, but not in the localisation task. These observations demonstrate that object selection and recognition are indeed separable stages in visual search, and that the activation of a sustained working memory representation of an object is not an automatic and inevitable consequence of its previous attentional selection. Even though they are dissociable, selective attention and working memory are usually closely linked. For example, the current content of working memory can affect the allocation of attention during visual search even when this content is irrelevant for the search task. When observers are asked to memorize a particular colour for subsequent recall before performing an independent visual search task, the presence of a distractor that matches the memorized colour impairs search performance (e.g., Downing,

2000; Soto, Heinke, Humphreys, & Blanco, 2005; Olivers, Meijer, & Theeuwes, 2005; Olivers & Eimer, 2011). This suggests that spatial attention can be biased towards memory-matching but currently task-irrelevant objects (see Olivers et al., 2011, for further discussion).

Sustaining a spatial focus of attention over time during the maintenance of visual object representations in working memory requires recurrent feedback from higher-order attentional control areas (e.g., Bundesen et al., 2005; Xu & Chun, 2009; Luck & Vogel, 2013; see also Bar, 2003; Hochstein & Ahissar, 2002, for the importance of recurrent feedback signals during object identification). In the absence of such recurrent feedback loops, any spatially specific enhancement of object processing that is triggered in visual cortex during the initial object selection stage is assumed to remain transient and fade rapidly, as was observed in the localisation task of our ERP study (Mazza et al., 2007). Regions in the intraparietal sulcus that are sensitive to working memory load and individual capacity limits (e.g., Todd & Marois, 2004; Xu & Chun, 2006) may play a central role in sustaining visual working memory representations of target objects in visual cortex when these objects have to be recognized. Why would object recognition require a spatially selective focus of attention that remains active for an extended period of time? Sustained focal attention may be needed to facilitate the binding of individual features within object representations in visual working memory (e.g., Wheeler & Treisman, 2002), in particular when these objects are no longer perceptually present. It may also be needed because object recognition processes involve comparisons between working memory representations of currently selected objects and stored object representations in long-term memory. Such comparison processes operate in real time, and may therefore require a sustained focus of spatial attention on particular working memory representations. More generally, recognition and identification processes are unlikely to be based exclusively on object representations in visual cortical maps, but will usually also involve interactions between these representations and other areas where semantic or episodic information about particular objects is stored (e.g., Sreenivasan et al., 2014). According to global neuronal workspace models of cognitive processing (e.g., Dehaene & Naccache, 2001), such long-distance interactions between different cortical regions are likely to be based on neural activation patterns that are maintained in a stable fashion over an extended period of time.

The research discussed in this section again highlights the continuous nature of the neural mechanisms that are active during visual search. Although cognitive models of visual attention emphasize the distinction between discrete selection and recognition stages, current ideas about the neural basis of attentional object selection and working memory maintenance suggest that these two stages are both based on spatially selective modulations of visual processing that differ primarily with respect to their temporal duration. Object selection is implemented by fast and transient spatially specific processing enhancements within cortical maps. Memory-based object recognition can take place when these spatial biases are actively sustained over time. Because time is a continuous variable, the transition from object selection to object recognition, that is, the transition from perceptual attention to working memory is also likely to be a continuous process.

## **6. Summary and Conclusions**

This article has reviewed the attentional processes that are responsible for our ability to find known target objects at unknown locations in visual search. These processes can be studied at the cognitive level and at the neural implementation level, and one aim of this review was to provide links between empirical findings and theoretical ideas at these two levels. The picture that has emerged from this discussion is quite different from traditional conceptualisations of selective attention that are based on the fundamental distinction between pre-attentive and attentive stages of perceptual processing. In such two-stage models, attentional selection mechanisms are located at the intersection between these two stages, and regulate the access of visual information to a central limited-capacity system. This two-stage architecture was proposed by Donald Broadbent in 1958, and has remained highly influential ever since. Indeed, as Jon Driver remarked in his review of attention research in the 20<sup>th</sup> century, Broadbent's ideas may "have been almost too influential; once exposed to them, it becomes hard to think about attentional issues in any other way" (Driver, 2001). The description of the attentional mechanisms involved in visual search that was developed in this review goes beyond such traditional two-stage models. It stresses the functional and temporal continuity of attentional processes and their neural basis, and questions the existence of a genuinely pre-attentive (i.e., goal-unselective)

processing stage in visual search. The idea that search goals are represented by attentional templates that are activated during the preparation for search (section 2), and the hypothesis that these preparatory processes produce spatially global goal-selective modulations during the subsequent attentional guidance phase (section 3) do not sit comfortably with a fundamental separation between pre-attentive and attentive processing and a distinct locus of attentional selection at the interface between these two stages.

Because attentional mechanisms are implemented by neural processes that unfold gradually in real time, they do not lend themselves easily to traditional discrete stage models of information processing. Nevertheless, it is still conceptually and heuristically useful to distinguish successive phases of attentional selectivity during visual search, such as the four phases described here (see Figure 1). Preparatory attentional templates can be set up prior to the arrival of visual input, and are reflected by goal-selective sustained baseline shifts of neural activity during the preparation for an upcoming attentional selection task (section 2). Attentional guidance processes start once visual input has been presented, and are based on goal-sensitive modulations of visual processing (feature-based attention) that operate in a spatially global fashion across the entire visual field (section 3). If preparatory attentional templates in visual areas represent search goals in a position-independent fashion, the transition from preparation to guidance is likely to be continuous rather than discrete. ‘Selection’ was defined as the emergence of spatially specific modulations of neural activity in visual cortical maps that facilitate the processing of task-relevant features or objects at particular locations (section 4). Again, the transition from spatially global feature-based attention to attentional object selection is best understood as a continuous process where spatially selective feature- and object-specific biases gradually develop across time. Finally, object recognition has been linked to the encoding of particular selected objects in visual working memory (section 5). Because working memory depends on spatially specific processing biases that were initially established during object selection and are then sustained over time, the transition between selection and maintenance is also continuous rather than discrete.

At the most general level, the attentional mechanisms described here can be characterized as the emergence and subsequent maintenance of spatially selective biases of visual processing in favour of potentially task-relevant objects. Such spatially focal biases are set up and sustained via recurrent interactions between visual cortex and higher-order



control areas, develop within the first 200 ms after visual input has been presented, and remain active until target objects have been successfully identified. The gradual transition from spatially global to spatially focused goal-selective neural activation patterns in cortical visual maps described here does not imply that this process will always eventually result in a single unitary focus of spatial attention. If working memory maintenance and the ability to track moving objects depend on spatial attention, the fact that multiple objects can be simultaneously tracked (e.g., Cavanagh & Alvarez, 2005) and held in visual working memory (e.g., Luck & Vogel, 1997; Cowan, 2001) suggests that multiple independent attentional foci can be maintained in parallel when required by current task demands. In other task contexts, a single focus of spatial attention may be required. This may be the case when complex visual objects such as individual faces have to be selected and maintained in working memory (e.g., Towler & Eimer, *in press*). Another example is the active exploration of visual scenes where the eyes move rapidly between different objects. In such situations, the selection of the next saccade target is based on allocating focal attention to one particular object in the visual field.

This review has described the attentional processes that are activated in visual search tasks where known targets have to be found at uncertain locations. In the absence of precise advance spatial information, the allocation of spatial attention to possible target objects will depend on spatially global attentional guidance processes. In tasks where the location of an upcoming task-relevant object is specified in advance (e.g., Posner et al., 1980), this type of guidance is not required, because observers can prepare for a particular target location before the arrival of visual input, and attention can then be allocated rapidly to this location. Preparatory baseline shifts of neural activity at pre-specified target locations have indeed been observed in spatial cueing experiments prior to the presentation of visual input (e.g., Luck, Chelazzi, Hillyard, & Desimone, 1997). ERP studies of cued visual-spatial attention (e.g., Mangun & Hillyard, 1991; Eimer, 1994) have found enhancements of visual P1 components for stimuli at attended versus unattended locations that start around 90 ms after stimulus onset. This shows that when target locations are known in advance, spatially specific attentional selection processes are triggered very rapidly. In contrast, the N2pc component that marks the selection of a target object at a previously unknown location in visual search tasks emerges nearly 100 ms later. This onset difference between spatially selective attentional effects in spatial cueing and visual search experiments fits

perfectly with the suggestion that target selection has to be preceded by a spatially global attentional guidance process in visual search, whereas no such guidance is required in spatial cueing tasks where target locations are known in advance.

The account of the attentional processes involved in visual search lined out in this review can provide a general framework that helps to integrate the results of behavioural and neuroscience studies of selective attention with diverse experimental procedures and measurement techniques. However, it does not provide a complete description of the complexities of visual search. In difficult visual search tasks, the target object is unlikely to be found on the basis of a single attentional episode that involves guidance, selection, and object recognition. When a target is hard to find, several iterations of these processes will be required that are triggered by a mismatch between visual object representations currently activated in working memory and an attentional template of the search target. The efficiency of visual search varies greatly across tasks (Wolfe, 1998), and such differences may reflect the number of iterations of a search cycle that are required before the target has been successfully recognized.

Attentional guidance in visual search may also be more complex than described in this article. While spatially global guidance processes are important when target locations are unknown, search does not always operate without any prior information about the likely position of task-relevant objects. During the processing of real-world visual scenes, context-sensitive spatial expectations can play an important role for the allocation of spatial attention (e.g., Hollingworth, 2009). Such contextual spatial information can also affect the detection of targets in simple search displays (“contextual cueing”; see Chun, 2000). In such situations, attentional guidance mechanisms may not operate in a spatially global fashion across the entire visual field, but could be restricted to specific areas that are most likely to contain the target objects. This suggests that context-dependent spatial expectations can be explicit or implicit parts of preparatory attentional templates. In addition, the allocation of attention in visual search tasks is not always exclusively guided in a goal-directed fashion by representations of target-defining features, but can also be affected by the perceptual salience of visual objects (e.g., Itti & Koch, 2001). This is acknowledged in the concept of priority-driven attentional guidance (e.g., Wolfe, 2007; Horowitz, Wolfe, Alvarez, Cohen, & Kuzmova, 2009), which is determined jointly by top-down information about target-defining features, context-dependent spatial expectations, and bottom-up salience signals. There

may also be situations where search operates without goal-selective guidance. In these cases, the allocation of spatial attention to particular objects would be determined either randomly, or exclusively by bottom-up salience signals. Working memory representations of target objects would still be required, because the recognition of selected objects as targets or nontargets depends on their comparison with stored representations of search goals. The original version of Feature Integration Theory (Treisman & Gelade, 1980) describes such a scenario where search operates without guidance.

The localization and recognition of target objects among distractors in visual search is based on processes that unfold in real time. The attentional mechanisms that are involved in visual search and their neural basis have been studied for decades, and this review has described the outlines of a general processing model. In this model, 'attention' is not characterized as a distinct stage of visual processing or as a dedicated top-down control system. Selective attention in visual search refers to the gradual and goal-dependent emergence of spatially selective processing biases for particular object representations in cortical visual maps.

## References

- Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience*, MIT Press, Cambridge, Mass, 183-218.
- Andersen, S. K., Hillyard, S. A., & Müller, M. M. (2008). Attention facilitates multiple stimulus features in parallel in human visual cortex. *Current Biology*, 18, 1006-1009.
- Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *Journal of Neuroscience*, 31, 1128-1138.
- Anderson, D. E., Vogel, E. K., & Awh, E. (2013). A common discrete resource for visual working memory and visual search. *Psychological Science*, 24, 929-938.
- Awh, E., Anillo-Vento, L., & Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 12, 840-847.
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5, 119-126.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 780-790.
- Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, 139, 201-208.
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15, 600-609.
- Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308, 529-534.
- Broadbent, D. E. (1958). Perception and communication. Pergamon, London.
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological Review*, 112, 291-328.
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, 9, 349-354.
- Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, 2, 671-676.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918-2940.
- Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, 4, 170-178.
- Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, 49, 1407-1409.
- Chun, M. M., & Johnson, M. K. (2011). Memory: Enduring traces of perceptual and reflective attention. *Neuron*, 72, 520-535.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-114.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness:

- basic evidence and a workspace framework. *Cognition*, 79, 1-37.
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 362, 761-772.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, 178, 363-380.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827-1837.
- Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, 11, 467-473.
- Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *Journal of Neuroscience*, 28, 4183-4191.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, 92, 53-78.
- Driver, J., & Frith, C. (2000). Shifting baselines in attention research. *Nature Reviews Neuroscience*, 1, 147-148.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, 87, 272-300.
- Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *The Quarterly Journal of Experimental Psychology*, 59, 2-27.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433-458.
- Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, 7, 255-261.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, 369, 313-315.
- Eimer, M. (1994). "Sensory gating" as a mechanism for visuospatial orienting: electrophysiological evidence from trial-by-trial cuing experiments. *Perception & Psychophysics*, 55, 667-675.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, 99, 225-234.
- Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, 99, 225-234.
- Eimer, M. & Grubert, A. (2014a). Spatial attention can be allocated rapidly and in parallel to new visual objects. *Current Biology*, 24, 193-198.
- Eimer, M., & Grubert, A. (2014b). The gradual emergence of spatially selective target processing in visual search: From feature-specific to object-based attentional control. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1819-1831.
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience*, 29, 15258-15265.
- Fannon, S. P., Saron, C. D., & Mangun, G. R. (2007). Baseline shifts do not predict attentional modulation of target processing during feature-based visual attention. *Frontiers in Human Neuroscience*, 1, 7.

- Fecteau, J. H., & Munoz, D. P. (2006). Saliency, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, 10, 382-390.
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in Cognitive Sciences*, 17, 134-141.
- Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, 173, 652-654.
- Ghorashi, S., Enns, J. T., Klein, R. M., & Di Lollo, V. (2010). Spatial selection and target identification are separable processes in visual search. *Journal of Vision*, 10, 7.
- Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, 1080, 63-72.
- Girelli, M., & Luck, S. J. (1997). Are the same attentional mechanisms used to detect visual search targets defined by color, orientation, and motion? *Journal of Cognitive Neuroscience*, 9, 238-253.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391, 481-484.
- Gratton, G., Corballis, P. M., & Jain, S. (1997). Hemispheric organization of visual memories. *Journal of Cognitive Neuroscience*, 9, 92-104.
- Grubert, A., & Eimer, M. (2013). Qualitative differences in the guidance of attention during single-color and multiple-color visual search: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 1433-1442.
- Grubert, A., & Eimer, M. (2015). Rapid parallel attentional target selection in single-colour and multiple-colour visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 86-101.
- Harrison, S.A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458, 632-635.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7, 498-504.
- Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36, 791-804.
- Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, 17, 273-291.
- Hopf, J. M., Boelmans, K., Schoenfeld, M. A., Luck, S. J., & Heinze, H. J. (2004). Attention to features precedes attention to locations in visual search: Evidence from electromagnetic brain responses in humans. *The Journal of Neuroscience*, 24, 1822-1832.
- Hopf, J. M., Luck, S.J., Girelli, M., Hagner, T., Mangun, G.R., Scheich, H., et al. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, 10, 1233-1241.
- Hornak, J., Duncan, J., & Gaffan, D. (2002). The role of the vertical meridian in visual memory for objects. *Neuropsychologia*, 40, 1873-1880.
- Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics*, 69, 172-184.
- Horowitz, T. S., Wolfe, J. M., Alvarez, G. A., Cohen, M. A., & Kuzmova, Y. I. (2009). The speed of free will. *The Quarterly Journal of Experimental Psychology*, 62, 2262-2288.

- Houtkamp, R., & Roelfsema, P. R. (2009). Matching of visual input to only one item at any one time. *Psychological Research*, 73, 317–326.
- Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, 114, 599–631.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2, 194–203.
- James, W. (1890/1981). *The Principles of Psychology*. Volume 1. Cambridge, Mass.: Harvard University Press.
- Katus, T., Grubert, A., & Eimer, M. (2014). Electrophysiological evidence for a sensory recruitment model of somatosensory working memory. *Cerebral Cortex*. Online publication; doi: 10.1093/cercor/bhu153.
- Khayat, P. S., Spekreijse, H., & Roelfsema, P. R. (2006). Attention lights up new object representations before the old ones fade away. *Journal of Neuroscience*, 26, 138–142.
- Kiss, M., Grubert, A., & Eimer, M. (2013). Top-down task sets for combined features: Behavioural and electrophysiological evidence for two stages in attentional object selection. *Attention, Perception, & Psychophysics*, 75, 216–228.
- Kravitz, D. J., Kriegeskorte, N., & Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cerebral Cortex*, 20, 2916–2925.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17, 26–49.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77, 24–42.
- Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1000–1014.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17, 391–400.
- Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 1057.
- Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, 14, 744–751.
- Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, 29, 317–322.
- Mazza, V., & Caramazza, A. (2011). Temporal brain dynamics of multiple object processing: the flexibility of individuation. *PLoS One*, 6, e17453.
- Mazza, V., Turatto, M., Umiltà, C., & Eimer, M. (2007). Attentional selection and identification of visual objects are reflected by distinct electrophysiological responses. *Experimental Brain Research*, 181, 531–536.
- McMains, S. A., Fehd, H. M., Emmanouil, T. A., & Kastner, S. (2007). Mechanisms of feature- and space-based attention: response modulation and baseline increases. *Journal of Neurophysiology*, 98, 2110–2121.

- Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: effects of practice and target similarity. *Journal of Experimental Psychology: Applied*, 15, 125-139.
- Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, 421, 370-373.
- Moore, T., Armstrong, K. M., & Fallah, M. (2003). Visuomotor origins of covert spatial attention. *Neuron*, 40, 671-683.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784.
- Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, 57, 1-17.
- Nako, R., Wu, R., & Eimer, M. (2014). Rapid guidance of visual search by object categories. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 50-60.
- Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and Category-Based Attentional Control During Search for Real-World Objects: Can You Find the Pants Among the Pans? *Journal of Experimental Psychology: Human Perception and Performance*, 40, 1283-1288.
- Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 411-421.
- Olivers, C. N. L., & Eimer, M. (2011). On the difference between working memory and attentional set. *Neuropsychologia*, 49, 1553-1558.
- Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1243-1265.
- Olivers, C. N. L., Peters, J., Houtkamp, R. & Roelfsema, P.R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends in Cognitive Sciences*, 15, 327-334.
- Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology*, 426, 505-518.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, 460, 94-97.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, 108, 12125-12130.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160.
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23-38.
- Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience*, 24, 3917-3925.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, 5, 631-632.



- Schall, J. D. (2004). On the role of frontal eye field in guiding attention and saccades. *Vision Research*, 44, 1453-1467.
- Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, 55, 301-312.
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 248-261.
- Sreenivasan K.K., Curtis, C.E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, 18, 82-89.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78, 364-375.
- Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences*, 106, 19569-19574.
- Stroud, M. J., Menneer, T., Cave, K. R., Donnelly, N., & Rayner, K. (2011). Search for multiple targets of different colours: misguided eye movements reveal a reduction of colour selectivity. *Applied Cognitive Psychology*, 25, 971-982.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135, 77-99.
- Thompson, K. G., & Bichot, N. P. (2005). A visual salience map in the primate frontal eye field. *Progress in Brain Research*, 147, 249-262.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428, 751-754.
- Towler, J., & Eimer, M. (in press). The focus of spatial attention determines the number and precision of face representations in working memory. *Cerebral Cortex*.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, 40, 201-237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 459.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428, 748-751.
- Warden, M.R., & Miller, E.K. (2010) Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15801-15810
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131, 48-64.
- Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1, 202-238.
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, 9, 33-39.
- Wolfe, J.M. (2007). Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99-119). New York: Oxford.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495-501.

- Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 787.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91-95.
- Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, 13, 167-174.
- Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel–serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 244.
- Zhang, W., & Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience*, 12, 24-25.

## Figure Legends

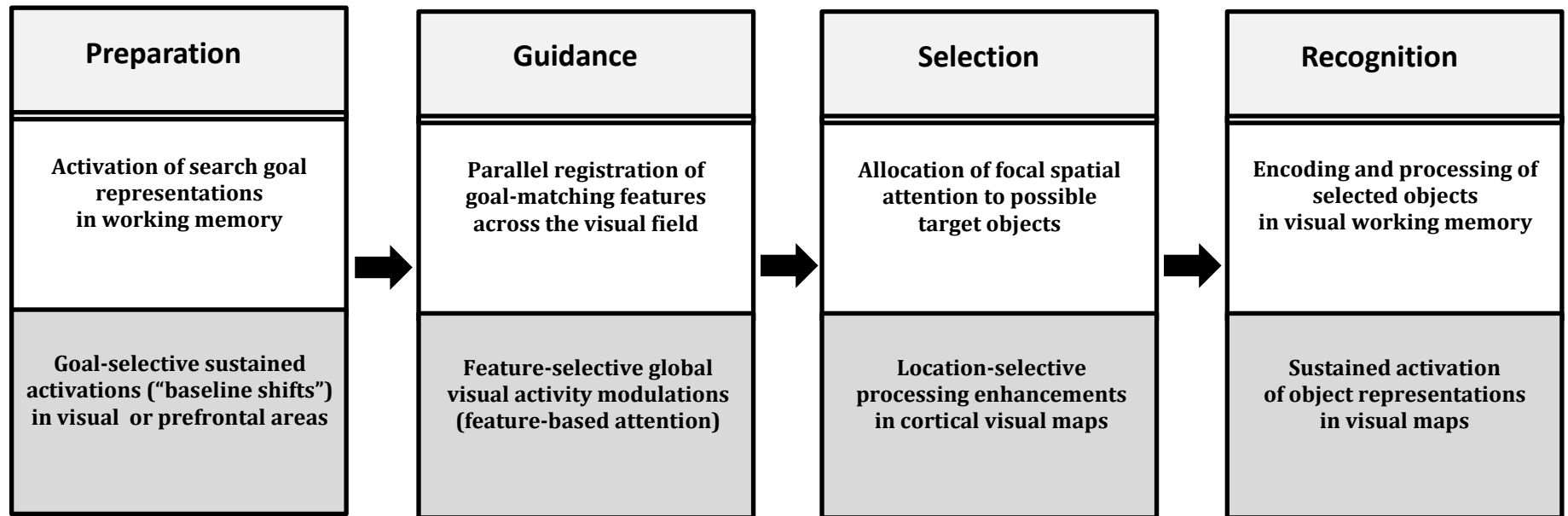
**Figure 1.** The four phases of attentional control during visual search discussed in this article. Preparation, guidance, selection, and recognition refer to different functions of selective attention that emerge at specific points in time during a search process. These functions are described at the cognitive level (white boxes), and in terms of their implementation at the neural level (dark grey boxes).

**Figure 2.** Top panel: Stimulus setup employed in the study by Eimer & Grubert (2014a). On each trial, two search displays were shown that contained a colour-defined target and a distractor on opposite sides. In the examples shown here, the red items are the targets. The two displays were presented in rapid succession, with a stimulus onset asynchrony (SOA) of 100 ms or 10 ms. One display contained two horizontal items and the other two vertical items. The horizontal target was equally likely to appear in display 1 or in display 2. Bottom panel: N2pc results in blocks where the SOA between the two displays was either 100 ms or 10 ms. ERPs at lateral posterior electrodes contralateral and ipsilateral to the horizontal target are shown together with N2pc difference waveforms obtained by subtracting ipsilateral from contralateral ERPs. All ERPs are plotted relative to the onset of display 1. With an SOA of 100 ms, the N2pc to a horizontal target in display 2 (H2) emerged 100 ms after the N2pc to a horizontal target in display 1 (H1). When the SOA was 10 ms, N2pc components to H1 and H2 targets were triggered within 10 ms of each other, and overlapped in time. This shows that the two targets were selected in parallel, with each selection process following its own independent time course. Data from Eimer & Grubert (2014a), reproduced in a different format.

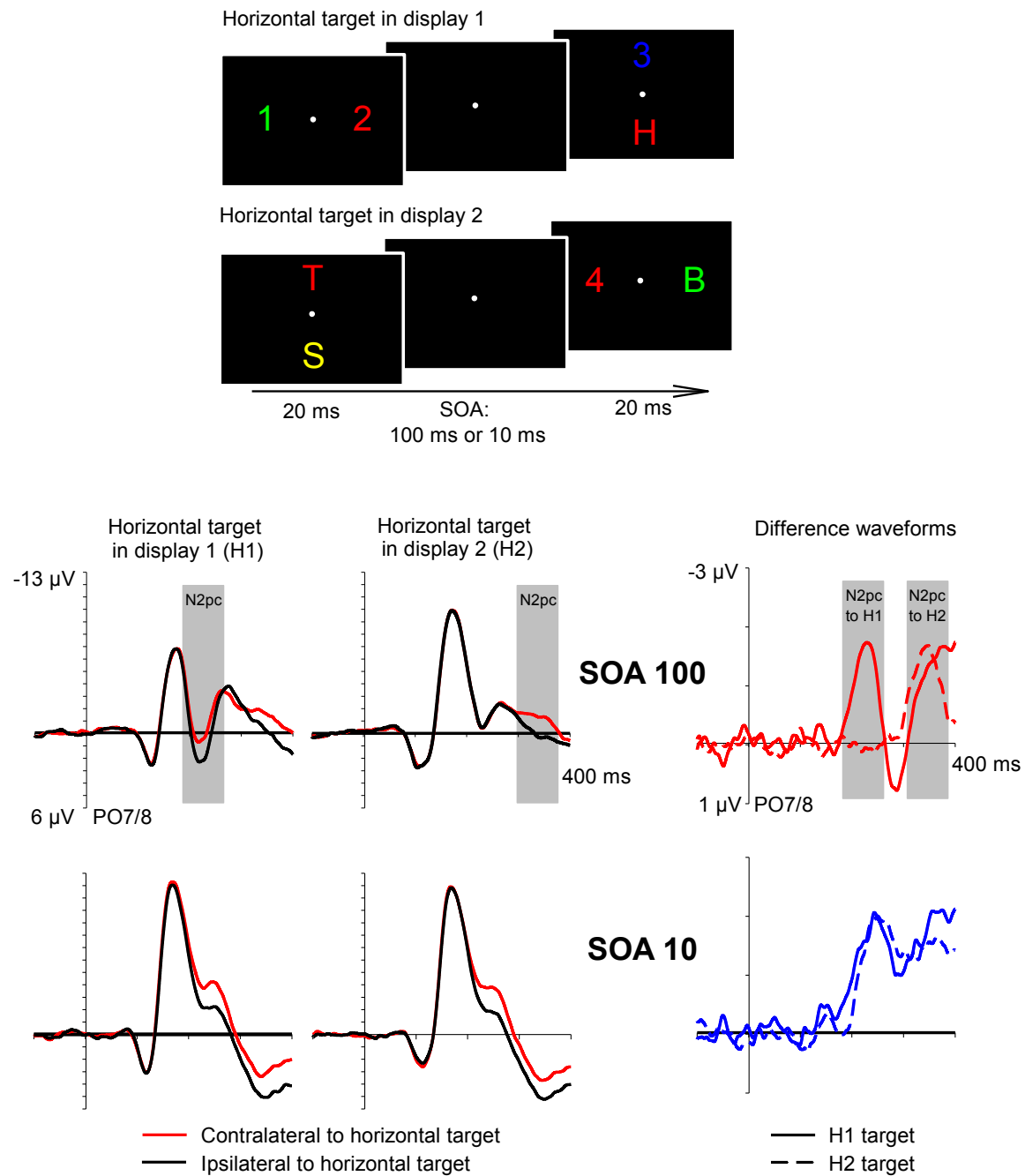
**Figure 3.** Left panel: Stimulus setup employed by Towler & Eimer (in press). Participants had to encode the two faces in a memory display, compare them to a centrally presented face in a test display, and decide whether one of the two memorized faces was repeated. The delay period between memory display offset and test display onset was very brief (200 ms). Right panel: ERPs measured at lateral posterior electrodes in response to memory displays in the 500 ms interval after display onset, for trials where one of the two faces in the memory displays was later repeated in the test display. On trials where this identity repetition was

detected correctly and rapidly, N2pc and CDA components were elicited contralateral to the face that would re-appear as the test face. On trials where participants failed to report the face repetition, N2pc and CDA components were elicited contralateral to the face in the memory display that was not repeated. These results show that attention was selectively allocated to one of the two faces in the memory displays. This focus of spatial attention determined which face was retained in working memory, and predicted the success or failure of face identity matching on individual trials. Data from Towler & Eimer (in press), reproduced in a different format.

Figure 1



## Figure 2



**Figure 3**

